

# Fisher Information Geometry for Shape Analysis

Angela De Sanctis and Stefano Antonio Gattone \*

University "G. d'Annunzio" of Chieti-Pescara, Pescara, Italy; a.desanctis@unich.it

\* Correspondence: gattone@unich.it

**Abstract:** The aim of this study is to model shapes from complex systems using Information Geometry tools. It is well-known that the Fisher information endows the statistical manifold, defined by a family of probability distributions, with a Riemannian metric, called the Fisher–Rao metric. With respect to this, geodesic paths are determined, minimizing information in Fisher sense. Under the hypothesis that it is possible to extract from the shape a finite number of representing points, called landmarks, we propose to model each of them with a probability distribution, as for example a multivariate Gaussian distribution. Then using the geodesic distance, induced by the Fisher–Rao metric, we can define a shape metric which enables us to quantify differences between shapes. The discriminative power of the proposed shape metric is tested performing a cluster analysis on the shapes of three different groups of specimens corresponding to three species of flatfish. Results show a better ability in recovering the true cluster structure with respect to other existing shape distances.

**Keywords:** landmarks; shape analysis; Fisher–Rao metric; information geometry; geodesics

---

## 1. Introduction

Shape analysis is a timely and interesting research field. Applications of shape analysis have been involved in various areas such as morphometry, image analysis, biology, database retrieval, and so on. In all these fields the aim could be to get an unsupervised classification of the objects in different clusters such that objects within a group are more similar in terms of shapes than they are in other groups. Then, the clustering of shapes is a longstanding challenge in the framework of geometric morphometrics [1], since the recognition of groups of similar morphologies, and then of the differences among these groups, is a key step of the analysis when geometric morphometrics protocols are applied. Shapes must be invariant to rotation, scale and translation so that a straightforward way to proceed is first to align the objects by using Procrustes analysis and then to apply standard clustering algorithms minimizing a given distance or dissimilarity measure evaluated within each cluster [2,3]. Similarly, [4] proposed to use a dissimilarity measure based on the inter-landmark distances and then apply standard statistical procedures such as hierarchical clustering or k-means clustering.

Under the hypothesis that it is possible to extract from the shape a finite number of representing points, called landmarks, we propose new statistical modeling of 2-dimensional shapes by representing each landmark by a bivariate Gaussian random variable. The means and the variances parameters becomes the coordinates of the statistical manifold. In particular, the variances reflect the uncertainty in the landmark's placement and the variability across a family of shapes. Within this framework, we derive a distance between two shapes using tools from Information Geometry [5,6], which considers statistical models as Riemannian manifolds with the Fisher–Rao metric. However, the associated geodesic distance has not yet been derived for the family of general multivariate normal distributions. Closed form expressions have been obtained only for isotropic and diagonal Gaussian distributions [7].

The paper is organized as follows: Section 2 recalls the main notions of Information Geometry. Section 3 reviews the statistical modeling of 2-dimensional shapes provided by Information

Geometry. In Section 4, a new shape distance derived from the Fisher–Rao metric is proposed. Its discriminative power is evaluated through a real application.

## 2. Information Geometry

Let  $I = [0, 1]$  and  $P$  a  $k$ -dimensional family of positive probability density functions

$$p : I \times \mathbf{R}^k \rightarrow \mathbf{R}^+$$

$(x, \theta) \mapsto p(x, \theta)$ , parametrized by  $\theta \in \mathbf{R}^k$ . In classical Information Geometry, the Fisher information matrix  $g$ , with generic  $(i, j)$ -entry

$$g_{ij}(\theta) = \int_0^1 p(x/\theta) \frac{\partial}{\partial \theta^i} \log p(x/\theta) \frac{\partial}{\partial \theta^j} \log p(x/\theta) dx. \quad (1)$$

is regarded as the most natural Riemannian structure on the parameter space [5,6].

From differential geometry we know that a metric matrix  $g$  defines an inner product on the tangent space of the manifold as follows:  $\langle u, v \rangle = u^T g_{ij} v$  with associated norm  $\|u\| = \sqrt{\langle u, u \rangle}$ .

Then the distance between two points  $P, Q$  of the manifold is given by the minimum of the lengths of all the piecewise smooth paths  $\gamma$  joining these two points. Precisely the length of a path is calculated by using the inner product:

$$\text{Length of } \gamma = \int_{\gamma} \|\gamma'(t)\| dt$$

and so

$$d(P, Q) = \min_{\gamma} \{\text{Length of } \gamma\}.$$

A curve that encompasses this shortest path is called a geodesic. In particular we consider the family of  $p$ -variate normal distributions:

$$p(x, \mu, \Sigma) = (2\pi)^{-\frac{p}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

where  $x = (x_1, x_2, \dots, x_p)^T$ ,  $\mu = (\mu_1, \mu_2, \dots, \mu_p)^T$  is the mean vector and  $\Sigma$  the covariance matrix. Note that the parameter space has dimension  $k = p + \frac{p(p+1)}{2}$ . We have three sub-cases [7]:

(i) Round Gaussian distributions:  $\Sigma = \sigma^2 I$

In this case the family can be identified with the  $(p + 1)$ -dimensional half space parameterized by  $(\mu_1, \mu_2, \dots, \mu_p, \sigma)$ ,  $\sigma > 0$ , and the Fisher information matrix is:

$$g(\mu_1, \mu_2, \dots, \mu_p, \sigma) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 & \dots & \dots & 0 & 0 \\ 0 & \frac{1}{\sigma^2} & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & \frac{1}{\sigma^2} & 0 \\ 0 & 0 & \dots & \dots & 0 & \frac{2p}{\sigma^2} \end{pmatrix}$$

Using a similarity transformation with the following matrix of the hyperbolic metric in the  $(p + 1)$ -dimensional half space

$$g(\mu_1, \mu_2, \dots, \mu_p, \sigma) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 & \dots & \dots & 0 & 0 \\ 0 & \frac{1}{\sigma^2} & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & \frac{1}{\sigma^2} & 0 \\ 0 & 0 & \dots & \dots & 0 & \frac{1}{\sigma^2} \end{pmatrix}$$

the closed form for the distance is given by

$$d((\bar{\mu}_1, \sigma_1), (\bar{\mu}_2, \sigma_2)) = \sqrt{2p} \ln \frac{|(\frac{\mu_1}{\sqrt{2p}}, \sigma_1) - (\frac{\mu_2}{\sqrt{2p}}, -\sigma_2)| + |(\frac{\mu_1}{\sqrt{2p}}, \sigma_1) - (\frac{\mu_2}{\sqrt{2p}}, \sigma_2)|}{|(\frac{\mu_1}{\sqrt{2p}}, \sigma_1) - (\frac{\mu_2}{\sqrt{2p}}, -\sigma_2)| - |(\frac{\mu_1}{\sqrt{2p}}, \sigma_1) - (\frac{\mu_2}{\sqrt{2p}}, \sigma_2)|} \quad (2)$$

where  $(\bar{\mu}_1, \sigma_1) = ((\mu_{11}, \dots, \mu_{1p}, \sigma_1)$  and  $(\bar{\mu}_2, \sigma_2) = ((\mu_{21}, \dots, \mu_{2p}, \sigma_2)$  besides  $|\cdot|$  is the usual Euclidean norm. The geodesics in the parameter space are contained in planes orthogonal to the hyperplane  $\sigma = 0$  and are either lines or half ellipses centered at this hyperplane. The curvature of the family is equal to  $\frac{-1}{p(p+1)}$ .

(ii) Diagonal Gaussian distributions:  $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)$

The family of all independent multivariate normal distributions is the intersection of half-spaces parameterized by  $(\mu_1, \sigma_1, \mu_2, \sigma_2, \dots, \mu_p, \sigma_p)$ ,  $\sigma_i > 0$ , so the Fisher information matrix is:

$$g((\mu_1, \sigma_1, \mu_2, \sigma_2, \dots, \mu_p, \sigma_p)) = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & \dots & 0 & 0 \\ 0 & \frac{2}{\sigma_1^2} & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & \frac{1}{\sigma_p^2} & 0 \\ 0 & 0 & \dots & \dots & 0 & \frac{2}{\sigma_p^2} \end{pmatrix}.$$

In this case the metric is a product metric, the curvature of the family is  $\frac{-1}{2(2p-1)}$ , and, using again the similarity with the hyperbolic metric, we have the following closed form for the distance

$$d((\mu_{11}, \sigma_{11}, \dots, \mu_{1p}, \sigma_{1p}), (\mu_{21}, \sigma_{21}, \dots, \mu_{2p}, \sigma_{2p})) = \sqrt{2 \sum_{i=1}^p \left( \ln \frac{|(\frac{\mu_{1i}}{\sqrt{2}}, \sigma_{1i}) - (\frac{\mu_{2i}}{\sqrt{2}}, -\sigma_{2i})| + |(\frac{\mu_{1i}}{\sqrt{2}}, \sigma_{1i}) - (\frac{\mu_{2i}}{\sqrt{2}}, \sigma_{2i})|}{|(\frac{\mu_{1i}}{\sqrt{2}}, \sigma_{1i}) - (\frac{\mu_{2i}}{\sqrt{2}}, -\sigma_{2i})| - |(\frac{\mu_{1i}}{\sqrt{2}}, \sigma_{1i}) - (\frac{\mu_{2i}}{\sqrt{2}}, \sigma_{2i})|} \right)^2} \quad (3)$$

(iii) General Gaussian distributions:  $\Sigma$  any symmetric positive definite covariance matrix

The analysis is much more difficult and it is not known a closed form for the associated distance.

### 3. Modeling of 2-Dimensional Shapes

We will consider only planar objects, as for example a section of the skull. The shape of the object consists of all information invariant under similarity transformations, that is translations, rotations and scaling [8]. Data from a shape are often realized as a set of points. Many methods allow us to extract a finite number of points, which are representative of the shape and are called landmarks. One way to compare shapes of different objects is to first register them on some common coordinate system for removing the similarity transformations [9,10]. Alternatively, Procrustes methods [11] may be used in which objects are scaled, rotated and translated so that their landmarks lie as close as possible to each other. Suppose we have a sample of  $n$  planar shapes. Let us denote the shape coordinates of the  $j$ -th configuration  $S_j$ ,  $j = 1, \dots, n$ , via its  $K$  landmarks  $\mu_j = \{\mu_{j1}, \mu_{j2}, \dots, \mu_{jK}\}$  with generic element  $\mu_{jk} = \{\mu_{jk1}, \mu_{jk2}\}$  for  $k = 1, \dots, K$ . For the  $k$ -th landmark an estimate of the coordinates covariance matrix  $\Sigma_k$  is given by

$$\Sigma_k = \frac{1}{n} \sum_{j=1}^n \text{vec}(\mu_{jk} - \bar{\mu}_k)(\text{vec}(\mu_{jk} - \bar{\mu}_k))^T \quad (4)$$

where  $\bar{\mu}_k$  denotes the  $k$ -th landmark coordinates of the mean shape  $\bar{\mu} = \frac{1}{n} \sum_{j=1}^n \mu_j$  and  $\text{vec}$  is the vectorization operator.

Our proposal is to model each landmark  $k$ , having shape coordinates  $\mu_k = \{\mu_{k1}, \mu_{k2}\}$ , with a bivariate Gaussian density. Assuming a round Gaussian distribution, case i) of section (2), we have  $\Sigma_k = \Sigma = \sigma^2 \mathbf{I}_2$  obtaining the following representation for the  $k$ -th landmark, for  $k = 1, \dots, K$ :

$$\frac{1}{2\pi\sigma^2} \exp\left\{-\frac{\|\mathbf{x} - \mu_k\|^2}{2\sigma^2}\right\} \quad (5)$$

where  $\mathbf{x}$  is a generic 2-dimensional vector. In the landmark representation (5),  $\sigma^2$  is a free parameter isotropic across all the  $K$  landmarks. Therefore, only the means are used as coordinates of the statistical manifold. We can relax the isotropic hypothesis assuming diagonal Gaussian distribution, case ii) of section (2). To this end, the new model for the representation of the  $k$ -th-landmark, for  $k = 1, \dots, K$ , is given by

$$\frac{1}{2\pi} |\Sigma_k|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mu_k)^T \Sigma_k^{-1} (\mathbf{x} - \mu_k)\right\} \quad (6)$$

where  $\Sigma_k = \sigma_k^2 \mathbf{I}_2$  and  $\sigma_k^2 = \{\sigma_{k1}^2, \sigma_{k2}^2\}$  is the vector containing the variances of the  $k$ -th landmark coordinates. Representation (6) allows us to express the  $k$ -th landmark coordinates as  $\theta_k = (\mu_k, \sigma_k)$  on a 4-dimensional manifold which is the Cartesian product of two half-planes.

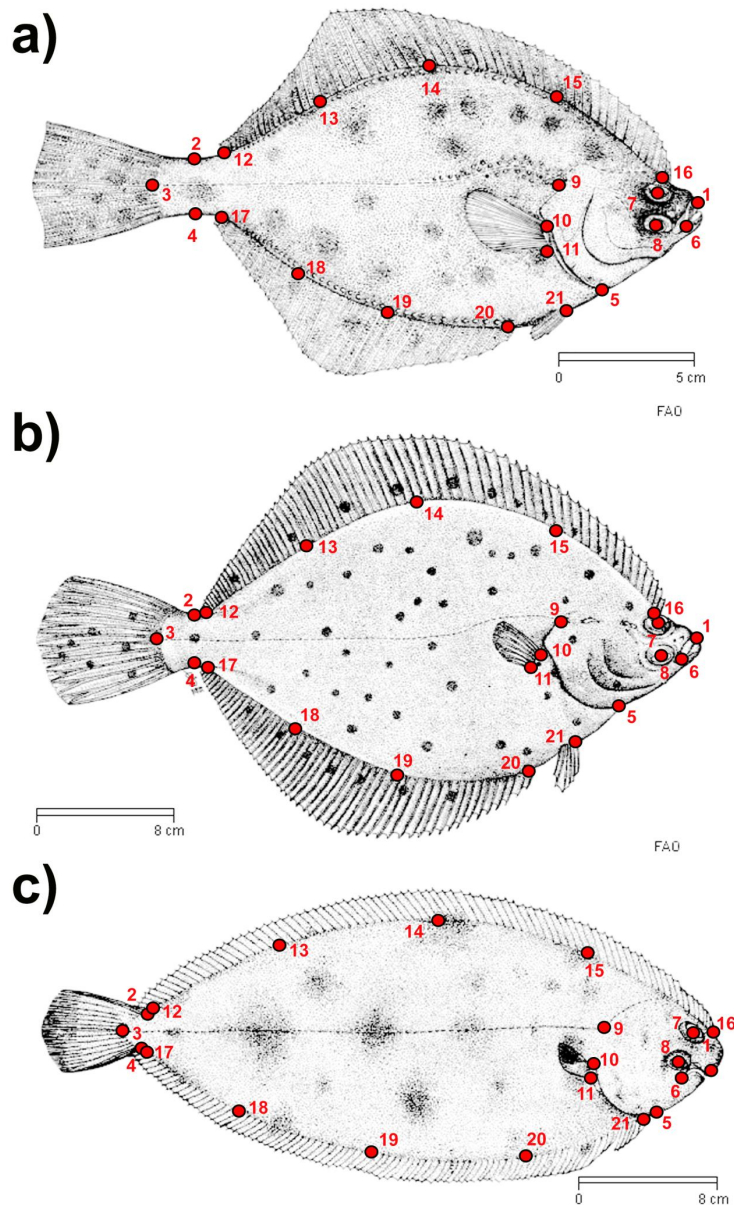
#### 4. Shape Metrics Based on Geodesic Distance

The landmarks representation as probability distributions enables to perform various type of analysis as for example quantifying the difference between shapes. Let  $S_1$  and  $S_2$  be two planar shapes. Denote the length of the geodesic path connecting the  $k$ -th landmarks of the two configurations by  $d(\theta_k^{S_1}, \theta_k^{S_2})$ . A shape metric for measuring the difference between  $S_1$  and  $S_2$  can be obtained evaluating the geodesic distances between the corresponding landmarks of the two configurations as follows:

$$d(S_1, S_2) = \sum_{k=1}^K d(\theta_k^{S_1}, \theta_k^{S_2}). \quad (7)$$

In Section 2 we provided closed form for the geodesic distance  $d(\theta_k^{S_1}, \theta_k^{S_2})$  according to the type of landmarks representation adopted. Under the isotropic variance assumption - landmarks representation (5) - the geodesic distance is computed with (2) while with the varying variance model - landmarks representation (6) - the geodesic distance is computed with (3). In order to apply (3) we need the covariance matrix in (6) to be diagonal. To this purpose we can apply an orthogonal transformation to the original landmarks coordinates which does not affect the analysis because it induces a rotation on the plane which leaves invariant the shape of a configuration.

The discriminative power of the proposed shape metrics is tested performing a cluster analysis on the shapes of three different groups of specimens corresponding to three species of flatfish. Namely, 60 individuals of plaice (*Pleuronectes platessa* L.) and 63 of flounder (*Platichthys flesus* L.) were collected in North Bull Island [12]. In addition, a group of 14 individuals of the species *S. solea* were used. These individuals were collected during the activities for the SOLEMON survey carried out the in the Adriatic Sea (Mediterranean Sea) during the autumn 2014. This last group of individuals would represent a kind of out-group with respect to the other two ones, since it belongs to a different and phylogenetically distant family (Soleidae) [13]. Each fish was photographed in lateral aspect and digital images were extensively described in [12]. In summary, the scheme of 21 landmarks described in Figure 1 was digitized for each individual of the three species.



**Figure 1.** Landmark configuration collected on (a) *Platichthys flesus*, (b) *Pleuronectes platessa*, and (c) *Solea solea*. 1, snout tip; 2, 3 and 4, points of maximum curvature of the peduncle; 5, insertion of the operculum on the lateral profile; 6, posterior extremity of premaxillar; 7 and 8, centres of the eyes; 9, beginning of the lateral line; 10 and 11, superior and inferior insertion of the pectoral fin; 12-16, semi-landmarks collected on dorsal fin; 17-20, semi-landmarks collected on anal fin; 21, insertion of first ray of the pelvic fin.

Three different shape metrics are computed:

- shape metric under the varying variance model (6) ( $dG$ )
- shape metric under the fixed variance model (5) ( $rG$ )
- Procrustes distance (PD)

The Procrustes distance is one of the main measures of difference between shapes. It is calculated by minimizing the Euclidean sum of squares between the landmark configurations using translation, rotation and scale. The computation of the geodesic distance within the model (5) requires the choice of the free parameter  $\sigma^2$ . In order to test the sensitiveness of the final clustering with respect to changes in the value of  $\sigma^2$ , we adopted three different values of  $\sigma^2$  given by the first ( $rG_{Q_1}$ ), the second ( $rG_{Q_2}$ ) and the third quartile ( $rG_{Q_3}$ ) of the values of the variances of each landmark and for each dimension. For each metric we computed the pairwise distances of all shapes. The obtained distance matrices are then used in a hierarchical clustering algorithm. The behaviour of the clustering has been evaluated by means of the a-Rand index [14]. Results for the solutions with three clusters are reported in Table 1 below.

**Table 1.** a-Rand index and number of miss-classified fishes.

Shape distance	$dG$	$rG_{Q_1}$	$rG_{Q_2}$	$rG_{Q_3}$	$PD$
aRand-index	0.8957	0.6669	0.7752	0.8708	0.3495
Number of missclassified fishes	4	14	9	5	35

Results show that the shape metric with the varying variance representation leads to the best performance in recovering the true cluster structure. The fixed variance model results in a worse cluster recovery with different behaviours depending on the value of the free parameter  $\sigma^2$ . For this particular data set, the Procrustes distance turns out to have a very poor performance in terms of cluster recovery.

As a conclusion, we only remark that the proposed shape representation allows also to study and predict the evolution in time of a shape [15].

**Acknowledgments:** We thank Tommaso Russo and Domitilla Pulcini for preparing the fish data and the support provided in commenting the results of the cluster analysis.

## References

1. Zelditch, M.L.; Swiderski, D.L.; Sheets, H.D.; Fink, W.L. *Geometric Morphometrics for Biologists: A Primer*; Elsevier Academic Press: London, UK, 2004.
2. Stoyan, D.; Stoyan, H. A further application of D.G. Kendall's Procrustes Analysis. *Biom. J.* **1990**, *32*, 293–301.
3. Srivastava, A.; Joshi, S.H.; Mio, W.; Liu, X. Statistical Shape analysis: Clustering, learning, and testing. *IEEE Trans. PAMI* **2005**, *27*, 590–602.
4. Lele, S.; Richtsmeier, J. *An Invariant Approach to Statistical Analysis of Shapes*; Chapman & Hall/CRC: New York, NY, USA, 2001.
5. Amari, S.; Nagaoka, H. *Methods of information geometry, volume 191 of Translations of Mathematical Monographs*; American Mathematical Society: Providence, RI, USA, 2000.
6. Murray, M.K.; Rice, J.W. *Differential Geometry and Statistics*; Chapman & Hall, New York, NY, USA, 1984.
7. Costa, S.; Santos, S.; Strapasson, J. Fisher information distance: A geometrical reading. *Discret. Appl. Math.* **2015**, *197*, 59–69.
8. Dryden, I.L.; Mardia, K.V. *Statistical Shape Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 1998.
9. Bookstein, F.L. Size and shape spaces for landmark data in two dimensions. *Stat. Sci.* **1986**, *1*, 181–242.
10. Kendall, D.G. Shape manifolds, Procrustean metrics and complex projective spaces. *Bull. Lond. Math. Soc.* **1984**, *16*, 81–121.
11. Goodall, C.R. Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc.* **1991**, *53*, 285–339.
12. Russo, T.; Pulcini, D.; O'Leary, A.; Cataudella, S.; Mariani, S. Relationship between body shape and trophic niche segregation in two closely related sympatric fishes. *J. Fish Biol.* **2008**, *73*, 809–828.
13. Berendzen, P.B.; Dimmick, W.W. Phylogenetic relationships of Pleuronectiformes based on molecular evidence. *Copeia* **2002**, *3*, 642–652.

14. Hubert, L.; Arabie, P. Comparing Partitions. *J. Classif.* **1985**, *2*, 193–218.
15. Sanctis, A.D.; Gattone, S. Methods of Information Geometry to model complex shapes. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 1271–1279.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).